

ID AF017250 standard; RNA; VRT; 5511 BP.
 XX
 AC AF017250;
 XX
 SV AF017250.1
 XX
 DT 06-JAN-1999 (Rel. 58, Created)
 DT 03-MAR-2000 (Rel. 62, Last updated, Version 2)
 XX
 DE Oreochromis aureus vitellogenin precursor (Vtg1) mRNA, complete cds.
 XX
 KW .
 XX
 OS Oreochromis aureus
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidei;
 OC Cichlidae; Oreochromis.
 XX
 RN [1]
 RP 1-5511
 RA Lim E.H., Lam T.J., Ding J.L.;
 RT "Cloning of Full-Length Oreochromis aureus Vitellogenin cDNA and its
 RT Deduced Primary Structure";
 RL Unpublished.
 XX
 RN [2]
 RP 1-5511
 RA Lim E.H., Lam T.J., Ding J.L.;
 RT ;
 RL Submitted (05-AUG-1997) to the EMBL/GenBank/DDBJ databases.
 RL School of Biological Sciences, National University of Singapore, 10 Kent
 RL Ridge Crescent 119260, Singapore
 XX
 DR SPTREMBL; Q9YGK0; Q9YGK0.
 XX

Key	Location/Qualifiers
<u>source</u>	1..5511
	/db_xref="taxon:47969"
	/sex="female"
	/organism="Oreochromis aureus"
	/tissue_type="liver"
	/dev_stage="estradiol-induced adult"
	/clone="pOaVtg1"
<u>5'UTR</u>	1..13
	/gene="Vtg1"
<u>sig_peptide</u>	14..61
	/gene="Vtg1"
<u>CDS</u>	14..5380
	/codon_start=1
	/db_xref="SPTREMBL:Q9YGK0"
	/gene="Vtg1"
	/function="egg yolk protein precursor"
	/product="vitellogenin precursor"
	/protein_id="AAD01615.1"
	/translation="MRVLVLALAVAGDQSNLAPGFASVKTYMYKYEAVLMGGLP GLARAGVKIRGKVLISATSANDYILKLVDLPQLLEYS GIWPKDPFHPATKLTALATQLS TPIKFYTNQVVGRLAAPPVSTTVLNIRGIINLLQLNVKKTQNVYEMQESGAHGVCK TNYVIREDARAERIHLTKTKDLNHCQEKIMKAIGLEHVEKCHDCEARGKSLKGTASNY

FT IMKPAPSGSLIMEAVAREVIEFSPFNILNGAAQMESKQILTFLDIENTPVDHARYTYVH
 FT RGSLLQYEHGSEILQTPIHLLRVTHAEAQIVSTLNHLVASNVAKVHEDAPLKFEVLIQVM
 FT RVARFETIESLWAQFKSRPDHRYWLLNAVPHIRTHAALKFLIEKLLANELSETEAAMAL
 FT LECLHSVTADQKTIELVRS LAENHRVKR NAVLNEIVMLGWTVISRFCKAQPS CSSDLV
 FT TPVHRQVAEAVETGDIDQLTVTLKCLDNAGHPASIKTIMKFLPGFGSAAARVPLKVQVD
 FT AVLALRRIAKREP KVMQEI AQAQLLMEKHLHAE LRMVAAMVLFETKLPVGLAASISTALI
 FT KEKNLQVVSFVYSYMKAMAKTTS PDHVSVA AACNVALRFLNPKLGR LNFYRSRAFHVDT
 FT YNNAWMMGAAASAVLINDAATVLP RMIMAKARTY MAGAYVDAFEVGV RTEGIQEALLKR
 FT RHENSENADRTIKIKQAMRALSEWRANPSSQALASMYVKVFGQEIAFANIDKSKVDQLI
 FT QFASGPLRNVRDAVNSVLSGYATHFAKPMLLGELRLILPTTVGLPMEISLITSAVTAA
 FT SVDVQATVSPPLPVNYRVSQ LLES DIQLRATVAPSLAMQTYAFMGVNTALIQAAVMTKA
 FT KVYTAVPAQIKARIDIVKGNLKVEFLSLQGIN TIASAHAETVAIARNVEDLPAARSTPL
 FT ISSETASQLSKASLNSKISRMAS SVTGGMSASSEIIPADLPSKIGRKM LPKTYRKKIR
 FT ASSRMLGFKAYAEIESHNAAYIRDCPLYALIGKHAASVRIAPASGPVIEKIEVEIQVGD
 FT KAAENMIKAIDMSEEEEALEDKNVLLKIKKILAPGLKNTTSSSSSSSSSSSSSSSSSSNKS
 FT SSSSSRSSSSQSSSSSRHSRSRKSQSSSSQSSRSPSSSSSSSSSSSSSSSSSSSRSS
 FT SRSSSSRSSSRRTKMADIVAPIITTSTRVSSSSSRASNSSSSSSASYLLSSSKRRSRSR
 FT SSSSSSSSSSSSSSSSSSSSSSSSKNSKRKSSNSKSSSSSRSSRRAQSKQQLLALKFRK
 FT NHVHRHAISTQRGSSHSARSFDSIYNKAKY LANTLTPAMSI AIRAVRVDHKVQGYQLA
 FT AYLDKQTNRLQLI FARVAEKDNWRICADIVQLSSHKLMAKTAWGAECKQYSTMIVAETG
 FT LLGHEPAARLKLTDWKLPGSIKHYAKRALKSIVPIAQEYGVNYAKAKNPRNQIKLTVAV
 FT ATETSMNIVLNTPKAIVYKRGVCLPVALPIGNTAAELQATRDNWADKMSYLVTKANAVE
 FT CSLINNTLTTFNNRKARDELPHSCYQVLAQDCTPELKFMVLLKKDQIQDQONQINVKISD
 FT IDVDMYRKNNAIAMVMNGVEIPNSNLPLYLHPSGNIHIRQSNEGITL NAPHSHGLQEVFLG
 FT FNELRVKVADWMKGKTCGACGTASGNVGDEYRTPSEQVTKDAISYAHSWVLSSNTCRDP
 FT SECSIKQESVKLEKRVIFEGVESKCYSEVPVLQCLPGCIPVRTTTVNVGFHCLPSDTTV
 FT DRSGLSFFFEKSIDLRDTAE AHLACRCTPQCA"

mat peptide

62..5377

/gene="Vtg1"

/product="vitellognein"

misc feature

3269..3946

/note="encodes putative phosvitin domain"

/gene="Vtg1"

3'UTR

5381..5511

/gene="Vtg1"

polyA signal

5465..5470

/gene="Vtg1"

XX

SQ

Sequence 5511 BP; 1474 A; 1443 C; 1299 G; 1295 T; 0 other;

acatccacca	gcatgaggg	tgcttgact	agctcttgct	gtggctctcg	cagtggggga	60
ccagtccaac	ttggcccag	gattgcctc	tgtaagacc	tacatgtaca	aatatgaagc	120
ggttcttatg	ggcggcctgc	ctgaagaggg	cctggctcga	gctgggggta	aaatccgggg	180
caaagttttg	atcagtgcaa	caagtgccaa	cgactacatt	ctgaagcttg	tagaccctca	240
gttgctggag	tacagtggca	tctggcccaa	agatcctttc	catccagcca	ccaagctcac	300
cacagccctg	gctactcagc	tctcgacacc	gatcaagttt	gagtatacaa	acggcggtgt	360
cgggagactg	gctgcacctc	ctgggggtctc	cacaacagtg	ctgaatatct	acaggggaat	420
catcaacctc	ctgcagctga	atgtaaagaa	gacacagaac	gtctacgaga	tgcaagagtc	480
tggagctcat	ggtgtgtgca	agaccaacta	tgtgatcagg	gaggacgcga	gagccgaacg	540
cattcatctg	accagagcca	aggacctgaa	ccactgccag	gagaaaatca	tgaaggccat	600
cggcttgga	cacgtagaga	aatgccatga	ttgtgaagct	agaggaaaga	gcctgaaggg	660
aactgcttcc	tataactaca	tcatgaagcc	agcaccagct	ggttctctga	ttatggaggc	720
tgctcgctaga	gaggctcatg	aattctcacc	tttcaacatt	ttgaatggcg	ctgctcagat	780
ggagtctaag	caaattctga	ccttcctgga	tatcgagaac	acccctgtgg	atcatgccag	840
atacacctat	gttcaccgcg	gatccctgca	gtatgagcat	ggcagcgaga	ttctccagac	900
acccatccat	cttctgaggg	tcacccatgc	cgaggctcag	attgtcagca	ctctgaacca	960
cctggtagcc	tccaacgtgg	ccaaggtcca	tgaagatgcc	cctctgaagt	ttgttgagct	1020
catccagggtg	atgcgtgtgg	ccagatttga	gactattgag	tcctctggg	ctcagtttaa	1080
atctagacct	gatcacaggt	actgggtact	gaatgctgtc	ccccacattc	gcactcacgc	1140
tgcgcttaag	ttcctcattg	agaagctcct	tgctaagtga	ttaagtgaga	ctgaagctgc	1200
tatggctctc	ttggaatgtc	tgcaactctgt	gacagctgac	cagaaaacca	ttgaacttgt	1260

cagaagcctg	gctgagaacc	acagagtga	acgtaacgct	gtgctcaacg	agattgtgat	1320
gctgggctgg	ggcactgtaa	tttccagggt	ctgtaaagcg	cagccatctt	gctcatctga	1380
tcttgtgaca	cctgtacata	gacaagttgc	agaggctgtt	gaaactgggt	acatcgatca	1440
gctcactgtc	acgctcaaat	gcctggataa	cgctggacat	cctgctagca	ttaagacaat	1500
catgaagttc	ctgcctggct	ttggcagtg	tgctgcccga	gtcccactca	aagttcaggt	1560
tgacgctgtt	ctagccctga	ggagaattgc	aaagagggaa	cccaagatgg	tccaggaaat	1620
agctgctcag	ttgctcatgg	aaaagcatct	ccatgcagaa	ctgcgtatgg	ttgctgccat	1680
ggtgctcttt	gagactaaac	tccccgtggg	tctagcagct	agcatttcca	cagccttgat	1740
caaagaaaag	aacctgcagg	tcgttagctt	tgtctactct	tacatgaagg	ccatggccaa	1800
gaccacatcc	cctgaccacg	tttctgttgc	tgcagcatgt	aatgttgcc	tgaggttcct	1860
caaccccaaa	ttaggcagac	tgaacttccg	ctacagccga	gccttccatg	tggataccta	1920
taacaatgcc	tgggatgatg	gtgctgccc	cagtgcctgc	ttaattaacg	acgctgcaac	1980
cgtgttacca	agaatgatta	tggccaaagc	ccgtacttac	atggccggag	cttatgttga	2040
tgctttttgag	gttggagtga	ggactgaggg	aatccaggag	gctcttttga	aaagacgaca	2100
tgaaaattct	gagaatgcag	acaggatcac	caagattaaa	caagccatga	gagctctttc	2160
tgagtggagg	gctaattcctt	cgagccaggc	cctggcctct	atgtatgtga	aggtcttcgg	2220
acaagaaatt	gcatttgcca	acattgacaa	atccaagggt	gaccagctta	tccagtttgc	2280
cagtggacct	ttgagaaacg	tattcagaga	tgctgtgaat	tctgtgctgt	ctggttatgc	2340
aacacatttt	gctaaaccaa	tgctgctcgg	tgagctccgt	ctcatccttc	ccaccactgt	2400
tgggttgccc	atggagatca	gcctcattac	atccgctgtg	actgctgcat	ctggtgacgt	2460
ccaagccact	gtgtcaccac	ctctgcctgt	caactaccga	gtttcccagc	ttctggagtc	2520
cgatatccaa	ctgagggcta	cagttgctcc	aagtcttgcc	atgcagacct	atgcattcat	2580
gggtgtgaac	accgccttaa	tccaggctgc	agtgatgaca	aaagccaaag	tttacacagc	2640
tgttcctgca	cagataaaaag	caaggattga	cattgttaag	ggcaacttga	aggttgagtt	2700
cctgtctctc	cagggcatta	acacaattgc	atctgcacat	gcggagacgg	ttgccattgc	2760
aagaaatgtg	gaagacctcc	cagccgcaag	aagcacacca	ctgatctcat	ctgaaactgc	2820
atcacaaactt	tcaaaggcct	ctctcaactc	aaagatctcc	aggatggcat	cctctgtgac	2880
tggtggcatg	tctgcgtcat	ctgaaatcat	tcctgctgac	ctgccaagta	agattgggag	2940
gaaaatgaaa	ctccctaaaa	cctacaggaa	gaaaatccgt	gcttcaagca	gaatgctagg	3000
attcaaggcc	tacgctgaga	ttgaatctca	caatgccgcg	tacatcagag	actgccctct	3060
ctacgctctg	atcggaaaagc	atgctgcttc	tgtaggatt	gctccagctt	ctggaccagt	3120
cattgagaag	attgaagtgtg	agattcaggt	cggagataaa	gcagcagaaa	atatgattaa	3180
agcgattgac	atgagcgaag	aggagggaagc	tcttgaggat	aagaatgtcc	tcttgaaaat	3240
caagaaaata	ctggcacctg	gtctcaagaa	caccacatca	tcttcctcca	gctcctccag	3300
ctcctcttca	tccagctcta	gtcccaacaa	gtctcttcca	tccagttccc	gctccagcag	3360
ctcccagtc	tccagctctc	gttcccatag	gtctcgctcc	agaaagtccc	agtctagcag	3420
ctctcagtc	agccgctctc	cctcaagctc	ttcctcctct	tctcctctct	catcatccag	3480
atcttcttcc	aggtcatctt	ccagatcatc	ttccagatct	tcttctaggt	cctcctctcg	3540
ctccagaact	aagatggctg	acattgttgc	tcctattatc	acgacgtcca	ccagagtgcg	3600
cagttcctcc	agtcgatcag	cctctaacag	ctcctccagc	agtgttcat	acttgctcag	3660
ctcatctaaa	agaagaagca	gaagcagaag	cagcagcagc	agtagcagca	gcagcagcag	3720
cagcagcagt	agcagcagca	gcagcagcag	cagtagcaag	aacagcaagc	gcagcaagag	3780
cagcaacagc	aagtcatcaa	gctctagggtc	ctctcggcgc	agtgtcaggt	ctaagcaaca	3840
actgcttgcc	ttgaagttca	gaaagaacca	cgtccacagg	catgccatct	ccacacagcg	3900
cggcagcagt	cacagcagtg	cccgcagctt	cgattccatc	tacaataagg	ccaagtacct	3960
cgctaacaca	ctcactcctg	ccatgtccat	tgcaatccgt	gccgtgagag	tcgaccacaa	4020
ggtccaggga	taccagctag	cagcttacct	ggacaaacag	accaatagac	tgagttgat	4080
ttttgccaga	gtcgctgaga	aggacaactg	gagaatctgt	gccgacattg	tgagctgag	4140
ttcgcacaag	ctgagggcca	agactgcctg	gggtgctgaa	tgcaagcaat	actccacct	4200
gattgtagct	gaaactgggtc	ttttgggtca	tgagcccgcg	gcccgttgga	agctgacctg	4260
ggacaaaactg	ccagggaagca	taaagcacta	gcgaagagg	gcgttgaaat	ccattgtccc	4320
tattgctcaa	gaatatggag	taaactacgc	aaaggccaag	aatcctcgta	atcaaatcaa	4380
actgactgta	gctgttgcta	ctgagacaag	catgaatatt	gtgctgaaca	caccaaaggc	4440
aatcgtttac	aagcgtgggg	tgtgtctacc	tgttgcctta	ccaattggaa	acactgctgc	4500
cgagctgcaa	gcgacccggg	acaactgggc	tgacaagatg	tcctatttgg	ttaccaaaagc	4560
taacgcagtt	gaatgctccc	tcatcaacaa	cacactgacc	acattcaaca	acaggaaagc	4620
tagagatgag	ctgccacact	cgtgctacca	ggctctggct	caggattgca	caccagaact	4680
caaattcatg	gttctgctga	agaaagacca	aatacaggat	cagaaccaga	tcaatgttaa	4740
gatttcagac	atcgatgtgg	acatgtatcg	gaagaacaac	gccattgcgg	tgatgggttaa	4800
cggagttgaa	atccctaaca	gcaacctgcc	atacctgcat	ccatcaggta	acatacatat	4860

aagacagtca	aatgaaggca	ttactctcaa	tgaccccagc	catggtcttc	aggaggtctt	4920
ccttggcttc	aacgagctga	gggttaaagt	tgacagactgg	atgaaaggaa	agacttggtg	4980
tgctgtgga	acggcaagcg	gaaatgtcgg	agacgagtac	cgcacacca	gtgaacaggt	5040
gaccaaggat	gccatcagct	acgcccactc	ctgggttctg	tcttcaaaca	cctgccgtga	5100
tccctccgag	tgttccatca	agcaggaatc	tgtgaagctg	gagaagcggg	tgatctttga	5160
aggtgtggag	tccaaatgct	actctgttga	gcccgtgctg	cagtgcctgc	ccggctgtat	5220
cccagtgaga	accactaccg	tcaacgttgg	ctttcactgc	ctgcccagtg	acacaactgt	5280
ggaccgttct	ggtctgagca	gcttctttga	gaagagcatc	gacctgaggg	atactgcaga	5340
agcccacctg	gcctgtcgct	gcactcctca	gtgtgcttaa	tctcattgcc	ttcggtcaca	5400
tgacattttg	ttttctgttt	taggcagggtg	ttgttaatta	gctttggctc	aacttgtaaa	5460
cttaaataaa	ttggaaaagc	atgtgaagca	aaaaaaaaa	aaaaaaaaa	a	5511

//

ID AF072686 standard; DNA; VRT; 11453 BP.
 XX
 AC AF072686; X93569; X93570; X93571; Z71336;
 XX
 SV AF072686.1
 XX
 DT 11-AUG-1999 (Rel. 60, Created)
 DT 30-NOV-1999 (Rel. 61, Last updated, Version 2)
 XX
 DE Oreochromis aureus vitellogenin gene, complete cds.
 XX
 KW .
 XX
 OS Oreochromis aureus
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidae;
 OC Cichlidae; Oreochromis.
 XX
 RN [1]
 RP 1-11453
 RX MEDLINE; 99145366.
 RA Teo B.Y., Tan N.S., Lim E.H., Lam T.J., Ding J.L.;
 RT "A novel piscine vitellogenin gene: structural and functional analyses of
 RT estrogen-inducible promoter";
 RL Mol. Cell. Endocrinol. 146(1-2):103-120(1998).
 XX
 RN [2]
 RP 1-11453
 RA Teo B.Y., Lim E.H., Ding J.L.;
 RT ;
 RL Submitted (18-JUN-1998) to the EMBL/GenBank/DDBJ databases.
 RL Department of Biological Sciences, National University of Singapore, 10
 RL Kent Ridge Crescent, Singapore 119260, Singapore
 XX
 DR SPTREMBL; Q9YGK0; Q9YGK0.
 XX
 CC On Nov 29, 1999 this sequence version replaced gi:1771394
 CC gi:1771393 gi:1771392 gi:1403509.
 XX

Key	Location/Qualifiers
<u>source</u>	1..11453
	/db_xref="taxon:47969"
	/sex="female"
	/organism="Oreochromis aureus"
	/clone="OaVg1"
	/tissue_type="blood"
	/dev_stage="adult"
<u>promoter</u>	1..1677
<u>protein_bind</u>	141..146
	/note="GATA binding site"
	/bound_moiety="GATA factor"
<u>protein_bind</u>	315..320
	/note="GATA binding site"
	/bound_moiety="GATA factor"
<u>protein_bind</u>	326..338
	/note="estrogen response element (ERE); imperfect ERE"
	/bound_moiety="estrogen receptor protein"
<u>protein_bind</u>	353..358

```

FT          /note="GATA binding site"
FT          /bound_moiety="GATA factor"
FT  protein_bind 440..445
FT          /note="GATA binding site"
FT          /bound_moiety="GATA factor"
FT  protein_bind 1001..1006
FT          /note="GATA binding site"
FT          /bound_moiety="GATA factor"
FT  protein_bind 1146..1158
FT          /note="estrogen response element (ERE); imperfect ERE"
FT          /bound_moiety="estrogen receptor protein"
FT  protein_bind 1200..1205
FT          /note="GATA binding site"
FT          /bound_moiety="GATA factor"
FT  TATA signal 1649..1655
FT  5'UTR      1678..1692
FT  mRNA      join(1678..1732,2065..2085,2193..2344,2441..2692,
FT              2793..2954,3044..3192,3280..3430,3519..3676,3755..3920,
FT              4000..4125,4211..4429,4538..4755,4850..4947,5035..5159,
FT              5288..5392,5502..5621,5709..5892,5990..6083,6177..6362,
FT              6462..6612,6694..6910,7017..7744,7834..7878,7957..8025,
FT              8117..8323,8407..8549,8736..8874,8978..9107,9224..9407,
FT              9497..9591,9686..9780,9906..10071,10213..10374,
FT              10471..10690)
FT          /product="vitellogenin"
FT  CDS      join(1693..1732,2065..2085,2193..2344,2441..2692,
FT              2793..2954,3044..3192,3280..3430,3519..3676,3755..3920,
FT              4000..4125,4211..4429,4538..4755,4850..4947,5035..5159,
FT              5288..5392,5502..5621,5709..5892,5990..6083,6177..6362,
FT              6462..6612,6694..6910,7017..7744,7834..7878,7957..8025,
FT              8117..8323,8407..8549,8736..8874,8978..9107,9224..9407,
FT              9497..9591,9686..9780,9906..10071,10213..10374,
FT              10471..10580)
FT          /codon_start=1
FT          /db_xref="SPTREMBL:Q9YGK0"
FT          /note="phospholipoglycoprotein; yolk protein precursor"
FT          /product="vitellogenin"
FT          /protein_id="AAD48085.1"
FT          /translation="MRVLVLALAVAVLAVGQSNLAPGFASVKTYMYKYEAVLMGGLPEE
FT          GLARAGVKIRGKVLISATSANDYILKLVDLPQLLEYSGIWPKDPFHPATKLTTALATQLS
FT          TPIKFEYTNQVVGRLAAPPVSTTVLNIYRGIINLLQLNVKKTQNVYEMQESGAHGVCK
FT          TNYVIREDAERAERIHLLTKTDLNHCQEKIMKAIGLEHVEKCHDCEARGKSLKGTASYN
FT          IMKPAPSGSLIMEAVAREVIEFSPFNILNGAAQMESKQILTFLDIENTPVDHARYTYVH
FT          RGSLOQYEHGSEILQTPIHLLRVTHAEAQIVSTLNLHLVASNVAKVHEDAPLKFVELIQVM
FT          RVARFETIESLWAQFKSRPDHRYWLLNAVPHIRTHAALKFLIEKLLANELSETEAAMAL
FT          LECLHSVTADQKTIELVRSLEAENHRVKNVAVLNEIVMLGWTVISRFCKAQPSCSSDLV
FT          TPVHRQVAEAVETGDDIDQLTVTLKCLDNAGHPASIKTIMKFLPGFGSAAARVPLKVQVD
FT          AVLALRRRIAKREPQMVEIAAQLLMEKHLHAEARMVAMVLFETKLPVGLAASISTALI
FT          KEKNLQVVSFVYSYMKAMAKTTSPDHVSVAACNVALRFLNPKLGRNLNFRYSRAFHVD
FT          YNNAWMMGAAASAVLINDAATVLPRLMIMAKARTYMAGAYVDAFEVGVRTGEGIEALLKR
FT          RHENSENADRITKIKQAMRALSEWRANPSSQALASMYVKVFGQEI AFANIDKSKVDQLI
FT          QFASGPLRNVFRDAVNSVLSGYATHFAKPMLLGELRLILPTTVGLPMEISLITSAVTAA
FT          SVDVQATVSPPLPVNYRVSQQLLESDIQLRATVAPSLAMQTYAFMGVNTALIQAAVMTKA
FT          KVYTAVPAQIKARIDIVKGNLKVFLSLQGINTIASAHAEVTAIARNVEDLPAARSTPL
FT          ISSETASQLSKASLNSKISRMASVTGGMSASSEIIPADLPSKIGRKMMLPKTYRKKIR
FT          ASSRMLGFKAYAEIESHNAAYIRDCPLYALIGKHAASVRIAPASGPVIEKIEVEIQVGD
FT          KAAENMIKAIDMSEEEEALEDKNVLLKIKKILAPGLKNTTSSSSSSSSSSSSSSSSSSNKS
FT          SSSSSRSSSSQSSSSSRHSRSRKSQSSSSQSSRSPSSSSSSSSSSSSSSSSSSSSSSSS
FT          SRSSSSRSSSRRTKMADIVAPIITTSTRVSSSSSRSSASNSSSSSSASYLLSSSKRRSR
FT          SSSSSSSSSSSSSSSSSSSSSSSSSSSSSSKNSKRSKSSNSKSSSSSRSSRRSAQSKQQLLALKFRK

```

FT NHVHRHAISTQRGSSSHSSARSFDSIYNKAKYLANTLTPAMSI AIRAVRVDHKVQGYQLA
 FT AYLDKQTNRLQLI FARVAEKDNWRICADIVQLSSHKLMAKTAWGAECKQYSTMIVAETG
 FT LLGHEPAARLKLTDWDLPGSIKHYAKRALKSIVPIAQEYGVNYAKAKNPRNQIKLTVAV
 FT ATETSMNIVLNTPKAIVYKRGVCLPVALPIGNTAAELQATRDNWADKMSYLVTKANAVE
 FT CSLINNTLTTFNNRKARDELPHSCYQVLAQDCTPELKFMVLLKKDQIQDQONQINVKISD
 FT IDVDMYRKNNIAVMVNGVEIPNSNLPYLHPSGNIHIROSNEGITLNAPSHGLQEVFLG
 FT FNELRVKVADWMKGKTCGACGTASGNVGEYRTPSEQVTKDAISYAHSWVLSSNTCRDP
 FT SECSIKQESVKLEKRVIFEGVESKCYSEVPVLQCLPGCIPVRTTTVNVGFHCLPSDTTV
 FT DRSGLSFFFEKSIDLRTAEAHACRCTPQCA"
 FT protein bind 2072..2084
 FT /note="estrogen response element (ERE); imperfect ERE"
 FT /bound_moiety="estrogen receptor protein"
 FT 3'UTR 10581..10690
 FT polyA signal 10665..10670
 FT polyA site 10690
 FT polyA signal 10739..10744
 XX
 SQ Sequence 11453 BP; 3340 A; 2463 C; 2285 G; 3365 T; 0 other;
 ctgcaggaat tcatttttct ttggcagaca ttcagatcta attacagttt cattcagcat 60
 tgctgagcat cgttatttat tcatttattc tttatacttg tattttctca acagtataaa 120
 taaccccgta acacttgat tgataatttt gcacacagtt tgctcacaca aacgcacaca 180
 attaattatt aagttaaata caatcagagg tgttggggtg ggggatacca gattattgga 240
 acatgtgaat tgcttgtaga acttgattta aaatactata aacttattgt gtgtacctga 300
 aattagcctg tgtgtgataa tggctgggaa acgtgacctc gtccacattc actgatagag 360
 acaaatgaaa gaggtaccaa caaacaggac agggctcggt tacaccggct ggtatcagcc 420
 tcaggactca cactgaggat gataaggaat gtgatgtaga agaataagggtg tggggatttc 480
 attgtcaciaa tttgtgtctg atgctttagt acatgagcct tttccttttc cccattgtta 540
 ataaatatag tggtagataa agaaaatctg aattaaatat tccttgtttt gtattatctg 600
 caaaataact gtgcaattca caaaatcatg aagtgtgaat gtgtttcaat gaaacaattt 660
 acatctagat gagaaaacat agatcagtta ataagtagtt acaaaatgcc atggatacat 720
 gtggataagt agtgactcag tgagtcagga ccagtcacct gttgtgggtc gaattgtttt 780
 taaatttttt gttagtagat ttaaattgag gtcttttaatt tagttttatt gctttttgtt 840
 tcttttatgt atattattaa cctctgaaaa tgtgtgcatg aggttttata tctgttgatg 900
 ccttataagc tgttcctggc gtccaaatca accccaagaa aaacagaact cataatttcg 960
 atataaaaatt taatattaca catttgattt ttattaaatt agataatttc ttttatgatt 1020
 cttaaaaaaaa ctgtgacagt ggttgtaag tcaacatgtg ttggcttatt gacttttgaa 1080
 aggtgctctt ggtcatggcc agtcgcagga gcaagatgac atgactgtga aaaagtcccc 1140
 tgccagggca cactgacatt tcaactgaatt gctaagtga aactaatcat atcagcgtga 1200
 gataattatt tcaactttat ccaataattg ttctttaaag catctcacca aaaaatcaat 1260
 ggtttgagtt aaatcgctta ttgaaaacag tggacacata acactgcttc tacagctgaa 1320
 atattggtat gttaatattt ctattcattc ataaattaaa ataactatga aaatagcaaa 1380
 atgcttaaga ttagaaaaact gattgtgttg aaaatgcgtg aagtcctaata gaagaaaaac 1440
 atattcttgt aattgggaca ttgggaaatc tctggaaatt tgtgaaaggc aagagccagg 1500
 tcaatacaac ttgtcacatc tgtgatgttt tctattagtt tattttacttt gtcaatctgg 1560
 gcaatcaggg catggttaaca tcaacaccaa agacagccgc catttagtta atgatacatt 1620
 tgatgggcaa cgtcagcaaa aaatctgctt aaaaaggacg cctctgcctg cagatcctca 1680
 catccaccag ccatgagggt gcttgtagta gctcttgctg tggctctcgc aggtgagtag 1740
 gaattattta caactacatg ttcttaaatg taacttaaaa ttcactcttct gacaaaaaac 1800
 agccttttta aagaattaga taacaactaa gaataagaat ttcaatagac ttgcagctga 1860
 gaagaattgt tactgcctgt gctagcagtt taatgctgcc agaataaac aatgctaatt 1920
 gccaaatgtc ttccagttta agcacattct ggcaaaatag catgaatttt atataattta 1980
 atatcactat ccctttacaa attcgcttcc ctatttcttg tcaactgtga caccaaatta 2040
 acatttattc ttctcatgtt acagtggggg accagtccaa cttgggtatg ttgttgctct 2100
 gttacattaa acagaaaagg cactgtcagt taatagagaa agcttctccc tcacctaga 2160
 aatgcacttt gtttctttgt cttttatcac agcccagga ttgcctctg ttaagaccta 2220
 catgtacaaa tatgaagcgg ttcttatggg cggcctgcct gaagagggcc tggctcgagc 2280
 tggggttaaa atccggggca aagttttgat cagtgcacaa agtgccaacg actacattct 2340
 gaaggtacag atcacgaggc agaattaaaa tgaaacatta ttgaaaccga caacttgaaa 2400
 catagcttgt ttctctctct ttctccgata ctttttgtag cttgtagacc ctcagttgct 2460
 ggagtacagt ggcattctgc ccaaatatcc tttccatcca gccaccaagc tcaccacagc 2520

cctggctact	cagctctcga	caccgatcaa	gtttgagtat	acaaacggcg	ttgtcgggag	2580
actggctgca	cctcctgggg	tctccacaac	agtgcgtaat	atctacaggg	gaatcatcaa	2640
cctcctgcag	ctgaatgtaa	agaagacaca	gaacgtctac	gagatgcaag	aggtacaaaa	2700
aaactcgctt	tcaccaggtc	cttcatacaa	aatgcttgaa	cactgacaga	aaataacaac	2760
aatctatatt	ttgattgtat	tttgtgatct	agtctggagc	tcatgggtgtg	tgcaagacca	2820
actatgtgat	cagggaggac	gcgagagccg	aacgcattca	tctgaccaag	accaaggacc	2880
tgaaccactg	ccaggagaaa	atcatgaagg	ccatcggtct	ggaacacgta	gagaaatgcc	2940
atgattgtga	agctgtaaga	gcattttcta	gttgtttcat	aatctctctt	atagaagtct	3000
tcatgctggc	agtttaaa	tcgtgttggt	tttatcattg	cagagaggaa	agagcctgaa	3060
gggaactgct	tcctataact	acatcatgaa	gccagcacc	agtgggttctc	tgattatgga	3120
ggctgtcgct	agagagggtca	tcgaattctc	acctttcaac	attttgaatg	gcgctgctca	3180
gatggagtct	aagtatgtgg	tttaaaactaa	aataatatgt	ttgattacaa	gcttacaagc	3240
attgtgccac	attgattgaa	cgattatctc	tgctcacagg	caaattctga	ccttcctgga	3300
tatcgagaac	accctgtgg	atcatgccag	atacacctat	gttcaccgag	gatccctgca	3360
gtatgagcat	ggcagcgaga	ttctccagac	acccatccat	cttctgaggg	tcacccatgc	3420
cgaggctcag	gtactgcagc	atttccttgc	agaaagtctt	ttctgaattt	accctgtagc	3480
cacaaacctc	tacatctgta	ttgtttgttt	cctcacagat	tgtcagcact	ctgaaccacc	3540
tggtagcctc	caacgtggcc	aaggtccatg	aagatgcccc	tctgaagttt	gttgagctca	3600
tccagggtgat	gcgtgtggcc	agatttgaga	ctattgagtc	cctctgggct	cagtttaa	3660
ctagacctga	tcacaggtaa	gtgcagcttt	gattcagcaa	aggtccacga	aactaacatg	3720
acacggataa	tggcattgtg	ctcctctttc	ccagggtactg	gttactgaat	gctgtcccc	3780
acattcgcac	tcacgctgcg	cttaagttcc	tcattgagaa	gctccttgct	aatgagttaa	3840
gtgagactga	agctgctatg	gctctcttgg	aatgtctgca	ctctgtgaca	gctgaccaga	3900
aaaccattga	acttgtcaga	gtaagtcttt	ttttaagctt	tcctatata	ttgcagtgc	3960
atgaaatcag	ttttatattg	tgtcttcacc	acttttcaga	gcctggctga	gaaccacaga	4020
gtgaaacgta	acgctgtgct	caacgagatt	gtgatgctgg	gctggggcac	tgtaatttcc	4080
aggttctgta	aagcgcagcc	atcttgctca	tctgatcttg	tgacagtatg	tattatttga	4140
acatgtttta	ttctgtctaaa	gttcagtttt	gacatcagtg	actttaatgt	ttttttacat	4200
tcttgacacg	cctgtacata	gacaagttgc	agaggctgtt	gaaactgggtg	acatcgatca	4260
gctcactgtc	acgctcaaat	gcctggataa	cgctggacat	cctgctagca	ttaagacaat	4320
catgaaagttc	ctgcctggct	ttggcagtg	tgctgcccga	gtcccactca	aagttcaggt	4380
tgacgctgtt	ctagccctga	ggagaattgc	aaagagggaa	cccaagatgg	tgagactaca	4440
ttttctgatt	cctaccatca	ataacaataa	atatacttga	gacctatttt	ctgtgcagtg	4500
ttcagttggg	gctatagtaa	tcggtctttc	tcattaggtc	caggaaatag	ctgctcagtt	4560
gctcatggaa	aagcatctcc	atgcagaact	gcgtatggtt	gctgccatgg	tgctctttga	4620
gactaaactc	cccgtgggtc	tagcagctag	catttccaca	gccttgatca	aagaaaagaa	4680
cctgcaggtc	gttagctttg	tctactctta	catgaaggcc	atggccaaga	ccacatcccc	4740
tgaccacgtt	tctgtgtaag	aacctaaatg	attattatgt	ttctaagatg	ctttatatct	4800
atgcatgcta	gtctattgaa	tctcaaatat	ccccttggtc	tattttaagt	gctgcagcat	4860
gtaatgttgc	cttgagggtc	ctcaacccca	aattaggcag	actgaacttc	cgctacagcc	4920
gagccttcca	tgtggatacc	tataacagta	agaagatacc	agataaataa	cttgctgttg	4980
cacaccttag	atgtgtgcat	gtgggtgtgat	gatgtttatc	catttgccctg	ccagatgcct	5040
ggatgatggg	tgctgccgcc	agtgccgtct	taattaacga	cgctgcaacc	gtgttaccaa	5100
gaatgattat	ggccaaagcc	cgtacttaca	tggccggagc	ttatgttgat	gcttttgagg	5160
taagacgttt	aaatgttacg	taagtaacta	aaacagacac	tcagcaatac	aattcaatga	5220
cttcgatatg	taaacaatac	atttttcagt	atctgcgctc	gagtgcattc	aataacgtac	5280
atttcaggtt	ggagtgagga	ctgagggaat	ccaggaggct	cttttgaaaa	gacgacatga	5340
aaattctgag	aatgcagaca	ggatcaccaa	gattaaacaa	gccatgagag	ctgtaagttg	5400
aactactctg	gcagagtga	aatgggattc	tgaggcacat	ttcttcatg	agagctgttt	5460
tttcttttac	ttgtttttaa	cttaatttat	ttccataata	gctttctgag	tggagggtca	5520
atccttcgag	ccaggccctg	gcctctatgt	atgtgaaggt	cttcggacaa	gaaattgcat	5580
ttgccaacat	tgacaaatcc	aaggttgacc	agcttatcca	ggtacagttt	ctccacagc	5640
ctcatatggt	ttcttagtaa	tgatgcatta	tgtaagttg	aaactaataa	gcaactttgt	5700
ctatacagtt	tgccagtggg	cctttgagaa	acgtattcag	agatgctgtg	aattctgtgc	5760
tgtctggtta	tgcaacacat	tttgctaaac	caatgctgct	cggtagagctc	cgtctcatcc	5820
ttcccaccac	tggtgggttg	cccatggaga	tcagcctcat	tacatccgct	gtgactgctg	5880
catctgttga	cggtaagttt	cattcagaca	gtttacaagt	ttacagaaag	gaaaaaatat	5940
atggatattg	atatgaaaaa	tatatgtttt	tcttggtgtc	ttttgctagt	ccaagccact	6000
gtgtcaccac	ctctgcctgt	caactaccga	gtttcccagc	ttctggagtc	cgatatccaa	6060
ctgaggggcta	cagttgctcc	aaggtaacaca	acagcttgag	aagcttcttc	tctgcatttt	6120

attgaatgaa	ttcagagata	actcttaatg	aatgaatatt	ttctgtcttt	atgcagtctt	6180
gccatgcaga	cctatgcatt	catgggtgtg	aacaccgcct	taatccaggc	tgcatgtatg	6240
acaaaagcca	aagttttacac	agctgttcct	gcacagataa	aagcaaggat	tgacattggt	6300
aagggcaact	tgaaggttga	gttcctgtct	ctccagggca	ttaacacaat	tgcatctgca	6360
cagtatgtac	atgccttgtc	actaattgcc	atgcattttca	aagatttttg	gtccttttcg	6420
gcatctgttt	ttattgggtg	gggttttttt	ctctcataca	gtgcggagac	ggttgccatt	6480
gcaagaaatg	tggaagacct	cccagccgca	agaagcacac	cactgatctc	atctgaaact	6540
gcatcacaac	tttcaaaggc	ctctctcaac	tcaaagatct	ccaggatggc	atcctctgtg	6600
actggtggca	tggtaagtgt	ggttcactcg	cctcctatgt	tcaacattca	actgtttgaa	6660
ataaacatgg	cgatctctgc	tgtgggtttta	cagtctgctg	catctgaaat	cattcctgct	6720
gacctgccaa	gtaagattgg	gaggaaaatg	aaactcccta	aaacctacag	gaagaaaatc	6780
cgtgcttcaa	gcagaatgct	aggattcaag	gcctacgctg	agattgaatc	tcacaatgcc	6840
gcctacatca	gagactgccc	tctctacgct	ctgatcgga	agcatgctgc	ttctgttagg	6900
attgctccag	gcaagtgcaa	gtggtgaact	tgaagcttta	acctaacaaa	tttgcaaaac	6960
attgtggaaa	agtcctgttt	ttgtttttta	ttattattct	ttaatattata	ttccagcttc	7020
tggaccagtc	attgagaaga	ttgaagttga	gattcaggct	ggagataaag	cagcagaaaa	7080
tatgattaaa	gcgattgaca	tgagcgaaga	ggaggaagct	cttgaggata	agaatgtcct	7140
cttgaaaatc	aagaaaatac	tggcacctgg	tctcaagaac	accacatcat	cttctccag	7200
ctcctccagc	tctcttccat	ccagctctag	ctccaacaag	tcttcttcat	ccagttcccg	7260
ctccagcagc	tcccagtcac	ccagctctcg	ttcccatagg	tctcgctcca	gaaagtccca	7320
gtctagcagc	tctcagtcac	gccgctctcc	ctcaagctct	tctcctctct	cctcctcttc	7380
atcatccaga	tcttcttcca	ggatcatctt	cagatcatct	tccagatctt	cttctaggct	7440
ctcctctcgc	tccagaacta	agatggctga	cattgtttgt	cctattatca	cgacgtccac	7500
cagagtgagc	agttcctcca	gtcgatcagc	ctctaacagc	tctccagca	gtgcttcata	7560
cttgctcagc	tcatctaaaa	gaagaagcag	aagcagaagc	agcagcagca	gtagcagcag	7620
cagcagcagc	agcagcagta	gcagcagcag	cagcagcagc	agtagcaaga	acagcaagcg	7680
cagcaagagc	agcaacagca	agtcacaaag	ctctagggtc	tctcggcgca	gtgctcagtc	7740
taaggtgaat	catcctctct	tgcgagcatg	tagaagtgtg	tggaaatttt	tagaatgtgg	7800
agcacttatt	ttacaattct	tttcttatta	cagcaacaac	tgcttgctct	gaagttcaga	7860
aagaaccacg	tccacagggg	aattattagg	catcacatca	aagcacatca	tgtcttctgt	7920
ttaagcaggc	taaatgtgtg	gtttccttgt	ttgcagcatg	ccatctccac	acagcgcggc	7980
agcagtcaca	gcagtgcccg	cagcttcgat	tccatctaca	ataaggtgat	cagtcaaatg	8040
aagatattta	cgatgttacg	cattaaactg	cgaataaaca	taattattaa	tatcaacctc	8100
tggttttcac	tttttaggca	agtacctcgc	taacacactc	actcctgcca	tgtccattgc	8160
aatccgtgcc	gtgagagtcg	accacaaggt	ccagggatac	cagctagcag	cttacctgga	8220
caaacagacc	aatagactgc	agttgatttt	tgccagagtc	gctgagaagg	acaactggag	8280
aatctgtgcc	gacattgtgc	agctgagttc	gcacaagctg	atggtaaaca	tcatgaactt	8340
atgaggaagt	aaccaactaa	tattaggtat	gatgcttagg	atttaactca	ctgaaattta	8400
ttttaggcca	agactgcctg	gggtgctgaa	tgaagcaaat	actccaccat	gattgtagct	8460
gaaactgggc	ttttgggtca	tgagcccgca	gcccgcttga	agctgacctg	ggacaaactg	8520
ccaggaagca	taaagcacta	cgcaaagagg	taaacttacg	gactttttcc	gattcataca	8580
tataatcagt	gctttttatta	aaaagctttt	gcagacattt	tgcagatgga	agtgacagag	8640
aattattttc	ttttgacatc	ctctcgaaac	tgtcattttat	actgaaaact	gacaaaagct	8700
ttataaaaaca	ctcagtaacc	tttatgcttt	tctagggcgt	tgaatccat	tgtccctatt	8760
gctcaagaat	atggagtaaa	ctacgcaaag	gccaagaatc	ctcgtaatca	aatcaaactg	8820
actgtagctg	ttgctactga	gacaagcatg	aatattgtgc	tgaacacacc	aaaggtacag	8880
ttttgttgcc	ccttcttctg	tatgcactga	aatgttttgc	attgcacttt	gtcactcaga	8940
gtgtttctaa	ctactttctg	actgtttaat	attccaggca	atcgtttaca	agcggtgggt	9000
gtgtctacct	gttcttttac	caattggaaa	cactgtgccc	gagctgcaag	cgacccggga	9060
caactgggct	gacaagatgt	cctattttgt	taccaagact	aacgcaggta	aaaatatttt	9120
aattaacaaa	aaaacaacaa	caacaacaac	aaaaacaggc	ttcatgtttc	tgtgggtgat	9180
ctttcagagg	ctaacatgtg	atttcctttt	cgttccactc	cagttgaatg	ctccctcatc	9240
aacaacacac	tgaccacatt	caacaacagg	aaagctagag	atgagctgcc	acactcgtgc	9300
taccaggtct	tggctcagga	ttgcacacca	gaactcaaat	tcatggttct	gctgaagaaa	9360
gaccaaatac	aggatcagaa	ccagatcaat	gttaagattt	cagacatgta	agagcagcca	9420
ggattttatgt	ggaaaagaaa	taacggctcag	cctgtgagcc	aaaaaatact	tatgtatttg	9480
aactgcttct	caacagcgat	gtggacatgt	atcggaagaa	caacgccatt	gcggtgatgg	9540
ttaacggagt	tgaatccct	aacagcaacc	tgccatacct	gcatccatca	ggttttttgt	9600
tttggtttgt	taatttaaca	aaaatattga	actgccagat	tgtgccattt	gaaaccatgc	9660
ttatatatat	ttgcacttac	tttaggtaac	atacatataa	gacagtcaaa	tgaaggcatt	9720

actctcaatg	cacccagcca	tgggtcttcag	gaggtcttcc	ttggcttcaa	cgagctgagg	9780
gtaatacaat	aaatgctgat	aattcatata	tctaatacgt	atattcatta	ttttctgagg	9840
caatttccta	tttgtatacc	tgttttaatc	tacattgtac	taatatctga	acctttttat	9900
ttcagggttaa	agttgcagac	tggatgaaaag	gaaagacttg	tgggtgcctgt	ggaacggcaa	9960
gcggaaatgt	cggagacgag	taccgcacac	ccagtgaaca	ggtgaccaag	gatgccatca	10020
gctacgcca	ctcctgggtt	ctgtcttcaa	acacctgccg	tgatccctcc	ggtaagcctt	10080
tgcaatatatt	ccttgacctt	tctctaaaag	aacattgaca	gcaatgcaa	atgtcccaac	10140
acaaacaatg	aatatatttc	atcaaacaat	gcacacatac	aatacttacc	ttcaccttct	10200
acactcttcc	agagtgttcc	atcaagcagg	aatctgtgaa	gctggagaag	cgggtgatct	10260
ttgaagggtg	ggagtccaaa	tgctactctg	ttgagcccgt	gctgcagtgc	ctgcccggct	10320
gtatcccagt	gagaaccact	accgtcaacg	ttggctttca	ctgcctgccc	agtggtgagt	10380
tgcaagaaaag	ccgagcttct	ttggaaaaca	ttatctctgt	tttgggtatta	cgtgcagtct	10440
ccaatcacgt	ctcctttttc	actgttgacg	acacaactgt	ggaccgttct	ggtctgagca	10500
gcttctttga	gaagagcatc	gacctgaggg	atactgcaga	agcccacctg	gcctgtcgct	10560
gcactcctca	gtgtgcttaa	tctcattgcc	ttcggtcaca	tgacattttg	ttttctgttt	10620
taggcagggtg	ttgttaatta	gctttggctc	aacttgtaaa	cttaaataaa	ttggaaaagc	10680
atgtgaagca	tgaaactgtt	gtgctcatct	ttcattatat	ttccttattc	tttattataa	10740
taaaacataa	ataagatcaa	cttaatttgt	cccaaaagtt	agttacagaa	acaaaaatct	10800
gaaaaaaatg	caacatgaca	aacgtcagta	aggctcgagtc	tactcaaaaag	aagcaaaaact	10860
aaatagaaaag	attaaagaaa	agtaaaaataa	ataaaaataaa	tctaaatcaa	atgcaataaa	10920
atgttatata	caaagcttga	tataatacag	taaggagctc	aaaatattaa	aatcttaagg	10980
aaatgcaatt	ataaaaaaac	aataaaaaatg	taaataaaaag	tgattttaca	gagtaagatt	11040
aaactagata	aaaatggatt	actaaaaacaa	atctgctcat	cctcttcttg	tatgtacttg	11100
ttgcacattt	tgcagtaaca	tgaaatcttt	gtaaacacag	tagtttatta	aaattgcatt	11160
tgtaatacaa	cctaagcagt	ttaggaacat	ttatcacggt	acaagtgaaa	gagacccttt	11220
tactgcactc	cttgatttgt	agaagggttg	cttattcgaa	tccttataat	gttaatcaca	11280
cgggtcacat	ctgctgggac	aggatgctgc	ggctcgtctc	atcctgcaat	ctggaaatgt	11340
tcacctttcc	ttttgcagaa	ctgcacattt	acattttggt	cttacatgca	caccagtaag	11400
aaagggctgc	tttcagcttt	aattcagagg	gtttaacaaa	aatcttgcat	cag	11453

//